文章编号: 1000-6788(2009)04-0111-08

基于广义非线性工具变量法的综列单位根检验

杨继生, 王少平

(华中科技大学 经济学院, 武汉 430074)

摘 要 当截面个体之间显著相关时, 非线性工具变量法 (NIV) 综列单位根检验存在严重的分布扭曲. 该文基于似无关回归形式的可行广义最小二乘法和非线性工具变量估计方法, 提出了广义非线性工具变量法 (GNIV) 综列单位根检验, 以修正 NIV 检验的分布扭曲. 在存在综列单位根的原假设下, GNIV 检验统计量的极限分布为标准正态分布, 而在备选假设下则趋向于负无穷大. 仿真实验结果显示, 在截面个体之间显著相关时, GNIV 检验的有限样本性质显著优于 Chang 的 NIV 检验和 Pesaran 的 CIPS 检验.

关键词 综列单位根检验; 非线性工具变量; 可行广义最小二乘法; 截面相关

中图分类号 F064.1 **文献标志码** A

Panel unit root test based on generalized nonlinear instrument variables estimation

YANG Ji-sheng, WANG Shao-ping

(School of Economics, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract This paper proposes a unit root test for panel data with cross sectional dependence. The proposed test generalizes the nonlinear IV unit root test proposed by Chang, and hence can correct the size distortion of Chang's test when cross sectional dependence is moderate to high. The main idea is to eliminate the cross sectional dependence through covariance matrix weighting and then apply Chang's test to the weighted data. Under some sufficient conditions, the proposed test statistic has a limiting standard normal distribution under the null hypothesis and converges to negative infinity under the alternative hypothesis. The finite sample performance of the proposed test is evaluated through a small scale simulation study. The simulation results show that the proposed test compares favorably to the tests proposed by Chang and Pesaran in the presence of cross sectional dependence.

Keywords panel unit root test; nonlinear instrument variables; FGLS; cross sectional dependence

1 引言

基于综列数据 (Panel data) 对经济变量进行实证分析, 增加了估计和检验统计量的自由度, 提高分析结论的可靠性, 所以, 基于综列数据的分析成为近期经济学和管理学实证研究的热点. 综列单位根检验用于检

收稿日期: 2007-11-19
资助项目: 国家自然科学基金 (70571026); 教育部人文社会科学研究青年项目 (08JC790043); 华中科技大学人文社科青年基金 (2007006)
作者简介: 杨继生 (1970–), 博士, 副教授, 研究方向: 计量经济学.

验综列数据的平稳性, 是目前国内外计量经济学方法论研究的热点问题之一.

目前, 综列单位根检验的研究主要关注于消除截面个体之间相关性的影响, 例如基于非线性工具变量法的检验 [1]、基于主成分分析的检验 [2]、基于退化时间趋势的检验 [3] 等. 其中, Chang[1] 提出的 S_N 综列单位根检验, 因其统计量的极限分布为标准正态分布而得到广泛应用. 但是, Im 和 Pesarsan[4] 在考察 Chang 的仿真实验时发现, S_N 检验仅在截面个体之间不相关或弱相关时才是有效的.

我们的研究结果验证了 Im 和 Pesarsan[4] 的发现, 即在 Chang 的数据生成过程中, 截面个体之间是弱相关的. 我们依照 Chang 的程序, 生成各截面个体扰动项之间的协方差矩阵. 结果显示, 截面个体的相关系数是有正有负的. 在 1000 次实验中, 当截面个体数量分别为 15、25 时, 相关系数的最大绝对值分别为 0.477 和 0.274. 所以, Chang 的仿真实验结果是在截面个体之间弱相关时得到的, 并且随着截面个体数量的增加, 个体之间的相关性也越来越弱. 对相关系数平均为 0.18 左右的综列数据, Im 和 Pesarsan 应用 Chang 的 S_N 检验, 发现 S_N 统计量不服从标准正态分布. 所以说, 对于截面相关的综列数据, S_N 检验不是一种稳健的检验方法.

为了保证截面相关时综列单位根检验的稳健性, 并使检验统计量的极限分布为标准分布, 本文结合似无关回归 (SUR) 形式的可行广义最小二乘法 (FGLS) 和非线性工具变量 (NIV) 估计方法, 提出了广义非线性工具变量法 (GNIV) 综列单位根检验. 本文第一部分给出模型的设定及其所需要的假定条件; 第二部分给出广义非线性工具变量法检验的基本思想、统计量的构造及其极限分布; 第三部分基于 Monte-Carlo 仿真实验考察新检验统计量的有限样本性质, 并与现有的检验统计量进行比较; 最后是相应的研究结论.

2 模型及其假定

考虑如下的综列数据生成过程:

$$y_{it} = \alpha_i y_{i,t-1} + u_{it}, \qquad i = 1, \cdots, N; \quad t = 1, \cdots, T \tag{1}$$

其中 α_i 为自回归系数, 误差项 u_{it} 服从如下的 AR(p) 过程:

$$W^i(L)u_{it} = \varepsilon_{it}, \qquad W^i(z) = 1 - \sum_{k=1}^{p} \rho_{i,k} z^k \tag{2}$$

其中 L 为滞后算子, 动态调整阶数 p 对各截面单元是相同的, $\{\rho_{i,k}, k = 1, 2, \cdots, p; i = 1, 2, \cdots, N\}$ 是相应的调整系数. 待检原假设为存在综列单位根, 即对任意 $i = 1, \cdots, N$, 有 $H_0 : \alpha_i = 1$; 对应的备选假设为: 对部分 i, 有 $H_1 : \alpha_i < 1$.

以下的假定条件与 Chang 的假定相似, 并且在时间序列单位根检验的文献中也经常用到. 定义 $\varepsilon_t^N = (\varepsilon_{1t}, \cdots, \varepsilon_{Nt})'$:

假定 1 对任意 $|z| \leq 1$, 有 $W^i(z) \neq 0, i = 1, \cdots, N$.

假定 2 (i) $\varepsilon_t^N, t = 1, 2, \cdots, T$, 为独立同分布向量序列, 并且关于勒贝格测度 (Lebesgue measure) 绝对连续; (ii) ε_t^N 具有零均值, 协方差矩阵为 $\Sigma = (\sigma_{ij})$; (iii) 存在 $l > 4$ 使 ε_t^N 满足 $E\{|\varepsilon_t^N|^l\} < \infty$; 有特征函数 φ, 对某些 $\tau > 0$, $\lim\limits_{\lambda \to \infty} |\lambda|^\tau \varphi(\lambda) = 0$ 成立.

假定 1 保证 (2) 中的 $AR(p)$ 过程是可逆的, 假定 2 设定了误差项的分布特征. 定义:

$$y_i = \begin{pmatrix} y_{i,p+1} \\ \vdots \\ y_{i,T} \end{pmatrix}, \quad y_{i,-1} = \begin{pmatrix} y_{i,p} \\ \vdots \\ y_{i,T-1} \end{pmatrix}, \quad X_i = \begin{pmatrix} x'_{i,p+1} \\ \vdots \\ x'_{i,T} \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,p+1} \\ \vdots \\ \varepsilon_{i,T} \end{pmatrix}$$

其中 $x'_{it} = (\Delta y_{i,t-1}, \cdots, \Delta y_{i,t-p})$. 模型 (1) 的矩阵表示为:

$$y_i = y_{i,-1}\alpha_i + X_i\beta_i + \varepsilon_i \tag{3}$$

其中 β_i 为滞后向量的调整系数.

众所周知, 在单位根的原假设下, 基于模型 (3) 的 OLS 估计, 估计量 $\hat{\alpha}_i$ 的 t 比率 (t 检验统计量) 的分布是非对称的, 即不再服从 t 分布. 因此, Chang 提出以 $y_{i,t-1}$ 的工具变量 $F(y_{i,t-1})$ 进行工具变量法估计, 以替代 OLS 估计. 工具变量生成函数 F 满足假定 3.

假定 3　$F(x)$ 是规范可积的并且满足 $\int_{-\infty}^{\infty} xF(x)\mathrm{d}x \neq 0$.

基于假定条件 1–3, Chang 得到了如下的关键结论. 当 $T \to \infty$,

$$\sqrt{T}\sum_{t=1}^{T}[F(y_{i,t-1})][F(y_{j,t-1})] \longrightarrow 0 \tag{4}$$

该结论显示, 无论截面单元之间是否相关, 工具变量法估计量 $\hat{\alpha}_i$ 和 $\hat{\alpha}_j$ 是渐近不相关的, 进而检验统计量 $S_N = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} t_{\alpha_i}$ 的极限分布为标准正态分布, 其中 t_{α_i} 为工具变量估计量 $\hat{\alpha}_i$ 的 t 比率. Chang 在证明 (4) 的结论时, 所用工具变量为 $F(y_{i,t-1}) = y_{i,t-1}e^{-c_i|y_{i,t-1}|}$, 其中 c_i 为常数, 显然, 该工具变量可以满足假定 3. 但是, 在其仿真实验的设定中, Chang 将 c_i 设定为:

$$c_i = KT^{-1/2}s^{-1} \tag{5}$$

其中 $s^2 = T^{-1}\sum_{t=1}^{T}(\Delta y_{it})^2$. 然而, Im 和 Pesaran 发现, 当使用 (5) 的设定形式时, 式 (4) 的结论将不再成立, 即在截面相关时, $\hat{\alpha}_i$ 和 $\hat{\alpha}_j$ 的 t 比率将不再是渐近不相关的. 原因在于, 此时 $y_{i,t-1}$ 的折算因子 $e^{-\frac{K|y_{i,t-1}|}{s_i\sqrt{T}}}$ 是有界的, $y_{i,t-1}$ 与 $F(y_{i,t-1})$ 将具有相同的渐近收敛阶, 所以在 $y_{i,t-1}$ 和 $y_{j,t-1}$ 相关时, 以 $F(y_{i,t-1})$ 和 $F(y_{j,t-1})$ 作为工具变量, 所得到的 $\hat{\alpha}_i$ 和 $\hat{\alpha}_j$ 的 t 比率将不再是渐近不相关的. 此时, 在截面相关时, 不同截面个体间的 t 比率将不能直接相加, Chang 基于 NIV 估计的 S_N 检验统计量不再具有标准正态的极限分布.

尽管 Chang 的 NIV 估计不能很好地消除截面个体之间的相关, 但却能够克服 OLS 估计的不足, 使得个体 t 比率的极限分布为标准正态分布, 进而, 可以基于该 t 比率, 建立具有标准正态分布的综列单位根检验统计量. 我们所需要做的, 只不过是消除不同截面个体之间的相关性罢了.

3　广义非线性工具变量法检验

基于上述分析, 我们提出如下的两步法检验程序. 首先, 基于 SUR 的思想, 运用各截面个体之间的即期协方差矩阵对综列数据进行加权处理, 以消除截面个体之间的相关性. 进而, 对加权后的综列数据进行 NIV 检验.

第一步　根据即期协方差矩阵进行数据加权

对模型 (3), 在综列单位根和同质性原假设下, 即 $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 1$ 和 $\beta_1 = \beta_2 = \cdots = \beta_N$ 时, 我们可以对原始数据进行如下的正交变换.

以 Σ 表示截面个体误差项的即期协方差矩阵, Γ 表示满足 $\Gamma\Gamma' = \Sigma$ 的对称不可逆矩阵. 令 $\Lambda = (\Gamma^{-1} \otimes I_{T-p-1})$, 其中 I_{T-p-1} 为 $(T-p-1)\times(T-p-1)$ 的单位矩阵, 符号 $\otimes$ 表示矩阵的 Kronecker 积. 定义:

$$\begin{aligned}
&Y^* = \Lambda(vec(y_1,\cdots,y_N)) = vec(y_1^*,\cdots,y_N^*)\\
&Y_{-1}^* = \Lambda(vec(y_{1,-1},\cdots,y_{N,-1})) = vec(y_{1,-1}^*,\cdots,y_{N,-1}^*)\\
&\Delta Y_{-k}^* = \Lambda(vec(\Delta y_{1,-k},\cdots,y_{N,-k})) = vec(\Delta y_{1,-k}^*,\cdots;\Delta y_{N,-k}^*),\quad k=1,\cdots,p\\
&\varepsilon^* = \Lambda(vec(\varepsilon_1,\cdots,\varepsilon_N)) = vec(\varepsilon_1^*,\cdots,\varepsilon_N^*)\\
&X_i^* = (\Delta y_{i,-1}^*,\cdots,\Delta y_{i,-p}^*),
\end{aligned} \tag{6}$$

则模型 (3) 可以表述为:

$$y_i^* = y_{i,-1}^*\alpha_i + X_i^*\beta_i + \varepsilon_i^* \tag{7}$$

显然, 转换后的误差项退除了截面个体之间的相关性, 其协方差矩阵为单位矩阵, 即 $(\varepsilon_t^*) \sim iid(0, I_N)$.

如果 Γ 是已知的, 则可以直接对 (7) 进行 NIV 检验. 因为 Γ 未知, 我们需要首先对 (3) 进行 NIV 估计, 根据估计得到的残差, 计算 Σ 和 Γ 的估计值.

对 (3) 进行 NIV 估计, 得到参数的一致估计量:

$$\tilde{r}_i = \begin{bmatrix} \tilde{\alpha}_i \\ \tilde{\beta}_i \end{bmatrix} = \begin{pmatrix} F(y_{i,-1})'y_{i,-1} & F(y_{i,-1})'X_i \\ X_i'y_{i,-1} & X_i'X_i \end{pmatrix}^{-1} \begin{pmatrix} F(y_{i,-1})' \\ X_i' \end{pmatrix} y_i \tag{8}$$

其回归残差为:

$$\tilde{\varepsilon}_{it} = y_{it} - y_{i,t-1}\tilde{\alpha}_i - x_{it}\tilde{\beta}_i \tag{9}$$

令 $\tilde{\varepsilon}_t^N = (\tilde{\varepsilon}_{1t}, \tilde{\varepsilon}_{2t}, \cdots, \tilde{\varepsilon}_{Nt})'$, 则残差的即期协方差矩阵为 $\hat{\Sigma} = \frac{1}{T}\sum\limits_{t=1}^{T}\tilde{\varepsilon}_t^N\tilde{\varepsilon}_t^{N'}$, 通过分解 $\hat{\Sigma} = \hat{\Gamma}\hat{\Gamma}'$, 得到估计值 $\hat{\Gamma}$. 以 $\hat{\Gamma}$ 替代 Γ, 得到 (6) 中的估计值 $\hat{y}_i^*$, $\hat{y}_{i,-1}^*$ 和 $\hat{X}_i^*$.

第二步 对加权后的数据进行 NIV 回归

令 $F(\hat{y}_{i,t-1}^*) = \hat{y}_{i,t-1}^* e^{c_i|\hat{y}_{i,t-1}^*|}$、$Z_i = (F(\hat{y}_{i,t-1}^*), \hat{X}_i^*)$、$V_i = (\hat{y}_{i,t-1}^*, \hat{X}_i^*)$. (7) 的 NIV 估计量、也就是 (3) 的 GNIV 估计量为:

$$\hat{r}_i = \begin{bmatrix} \hat{\alpha}_i \\ \hat{\beta}_i \end{bmatrix} = (Z_i'V_i)^{-1}Z_i'\hat{y}_i^* \tag{10}$$

$$\mathrm{Var}(\hat{r}_i) = (Z_i'V_i)^{-1}(Z_i'Z_i)(V_iZ_i')^{-1} \tag{11}$$

以 $\hat{V}_{\alpha_i}$ 表示 $\hat{\alpha}_i$ 的方差估计量, 则

$$\hat{V}_{\alpha_i} = \hat{B}_i^{-2}\hat{C}_i \tag{12}$$

其中:

$$\hat{B}_i = F(\hat{y}_{i,-1}^*)'\hat{y}_{i,-1}^* - F(\hat{y}_{i,-1}^*)'\hat{X}_i^*(\hat{X}_i^{*\prime}\hat{X}_i^*)^{-1}\hat{X}_i^{*\prime}\hat{y}_{i,-1}^*$$

$$\hat{C}_i = F(\hat{y}_{i,-1}^*)'F(\hat{y}_{i,-1}^*) - F(\hat{y}_{i,-1}^*)'\hat{X}_i^*(\hat{X}_i^{*\prime}\hat{X}_i^*)^{-1}\hat{X}_i^{*\prime}F(\hat{y}_{i,-1}^*)$$

相应地, $\hat{\alpha}_i$ 的 t 比率为:

$$\hat{t}_{\alpha_i} = \frac{\hat{\alpha}_i - 1}{\sqrt{\hat{V}_{\alpha_i}}} \tag{13}$$

命题 1 给出了其极限分布特征.

命题 1 如果假定 1 – 3 成立, 在单位根原假设下, 当 $T \to \infty$, 对于任意的 $i = 1, \cdots, N$, $\hat{\alpha}_i$ 的 t 比率的极限分布为:

$$\hat{t}_{\alpha_i} \to_d N(0,1) \tag{14}$$

对 $i \neq j$, 不同截面 t 比率的相关系数为:

$$\mathrm{cor}(\hat{t}_{\alpha_i}, \hat{t}_{\alpha_j}) \to_p 0 \tag{15}$$

其中 cor 表示相关系数.[1]

由于协方差矩阵 $\hat{\Sigma} = \hat{\Gamma}\hat{\Gamma}'$ 是一致估计量, 命题 1 可以由 Chang 的结论直接得出, 但为了表述清楚, 本文附录中还是根据 Chang 的推导给出了简要的证明过程.

命题 1 表明, 尽管在综列数据 y_{it} 的生成过程中, 各截面个体是相关的, 但是 GNIV 估计量 $\hat{\alpha}_i$ 的 t 比率在各截面个体间却是渐近不相关的. 因此, 各截面个体的 t 比率可以直接加总, 构造原假设为综列单位根的 GNIV 检验统计量, 即:

$$\hat{S}_N^* = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\hat{t}_{\alpha_i}. \tag{16}$$

基于定理 1 的结论, 可以直接得出 GNIV 检验统计量的渐近性质.

推论 1 如果假定 1–3 成立, 在综列单位根和同质性的原假设下, 当 $T \to \infty$, GNIV 检验统计量的极限分布为:

$$\hat{S}_N^* \to_d N(0,1). \tag{17}$$

命题 1 和推论 1 的结论可以很容易地扩展到含有确定性趋势的综列数据中, 我们所需要做的就是对截矩和时间趋势进行退势处理, 而后对退除均值和时间趋势后的综列数据进行 GNIV 检验. So 和 Shin[7] 和 Elliott、Rothenberg 和 Stock[6] 的退势方法均可以用来进行退除均值和退除时间趋势的处理.

1. 当 $N = 2$ 时, 我们将 ε_1 和 ε_2 的相关系数设定为 0.8, 在原假设 $\alpha_1 = \alpha_2 = 1$ 下, 基于 1000 次重复的仿真实验结果, 计算得到的 $\hat{t}_1$ 和 $\hat{t}_2$ 的相关系数值为 0.04.

4 仿真实验

我们通过一组仿真实验来考察 GNIV 检验分别在截面单元之间不相关、弱相关、强相关下的有限样本性质和检验势, 并将其与 Chang 的 S_N 检验和 Pesaran[5] 的 CIPS 检验进行比较.

4.1 一般性截面相关设定下的数据生成过程

依照模型 (1) 生成综列数据 $\{y_{it}\}$, 其中 $\{u_{it}\}$ 被设定为如下的 $AR(1)$ 过程:

$$u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it} \tag{18}$$

新息 $\varepsilon_t = (\varepsilon_{1t}, \cdots, \varepsilon_{Nt})'$ 由均值为零协方差矩阵为 Σ 的 N 维正态分布随机生成. 协方差矩阵 $\Sigma = (\sigma_{ij})$ 的主对角线元素固定为 $\sigma_{ii} = 1$, $i = 1, \cdots, N$, 其他元素在截面单元之间强相关、弱相关下分别设定为 $\sigma_{ij} = 0.8$ 和 $\sigma_{ij} = 0.3$, $i \neq j$. 为了考察截面不相关时各检验的表现, 我们同时给出了 $\sigma_{ij} = 0(i \neq j)$ 时的仿真结果. 自回归系数 ρ_i 在原假设下被设定为 0.3, 在备选假设下由均匀分布 $U[0.2, 0.4]$ 随机生成.

在原假设下, 考察统计量在有限样本下的检验水平 (Finite sample size) 时, 令 $\alpha_i = 1$, $i = 1, \cdots, N$. 当在备选假设下考察检验势 (Test power) 时, α_i 由均匀分布 $U[0.85, 0.99]$ 随机生成.

根据仿真实验的结果, 我们发现如下的 NIV 生成函数在有限样本下表现良好:

$$F(\hat{y}^*_{i,t-1}) = \hat{y}^*_{i,t-1} e^{-c_i |\hat{y}^*_{i,t-1}|}, \text{其中} c_i = 2.2N^{-1/4}T^{-1/2}s^{-1}, \quad s^2 = T^{-1}\sum_{t=1}^{T}(\Delta y^*_{it})^2 \tag{19}$$

表 1 和表 2 分别给出了 GNIV 检验在有限样本下的检验水平和检验势、以及与 S_N 检验的比较, 显著性水平分别选取为 0.01、0.05 和 0.10.

表 1　检验水平 (一般性截面相关设定, $\rho_i = 0.3, \alpha_i = 1$)

N		T=50			T=100		
		1%	5%	10%	1%	5%	10%
$\sigma_{ij} = 0.8$							
5	S_N^*	0.010	0.057	0.107	0.021	0.060	0.108
	S_N	0.040	0.158	0.217	0.055	0.181	0.233
15	S_N^*	0.014	0.056	0.088	0.020	0.054	0.098
	S_N	0.182	0.290	0.341	0.178	0.292	0.344
25	S_N^*	0.009	0.035	0.064	0.014	0.057	0.096
	S_N	0.256	0.371	0.402	0.250	0.355	0.392
$\sigma_{ij} = 0.3$							
5	S_N^*	0.014	0.064	0.117	0.019	0.070	0.134
	S_N	0.010	0.066	0.111	0.011	0.070	0.117
15	S_N^*	0.015	0.054	0.092	0.014	0.068	0.112
	S_N	0.012	0.077	0.122	0.011	0.080	0.120
25	S_N^*	0.008	0.036	0.072	0.019	0.060	0.107
	S_N	0.018	0.095	0.145	0.023	0.095	0.140
$\sigma_{ij} = 0$							
5	S_N^*	0.016	0.066	0.115	0.019	0.068	0.112
	S_N	0.018	0.055	0.094	0.016	0.056	0.089
15	S_N^*	0.016	0.056	0.100	0.017	0.076	0.133
	S_N	0.013	0.047	0.089	0.012	0.048	0.095
25	S_N^*	0.013	0.036	0.060	0.015	0.056	0.100
	S_N	0.010	0.038	0.078	0.010	0.038	0.073

表 2 检验势 (一般性截面相关设定, $\rho_i \sim U[0.2, 0.4], \alpha_i \sim U[0.85, 0.99]$)

N		T=50 1%	T=50 5%	T=50 10%	T=100 1%	T=100 5%	T=100 10%
$\sigma_{ij} = 0.8$							
5	S_N^*	0.433	0.692	0.783	0.833	0.925	0.953
	S_N	0.531	0.704	0.785	0.857	0.943	0.962
15	S_N^*	0.822	0.894	0.925	0.983	0.993	0.998
	S_N	0.807	0.866	0.894	0.973	0.990	0.992
25	S_N^*	0.817	0.888	0.911	0.995	1	1
	S_N	0.833	0.886	0.910	0.987	0.995	0.996
$\sigma_{ij} = 0.3$							
5	S_N^*	0.599	0.825	0.898	0.940	0.984	0.992
	S_N	0.566	0.780	0.875	0.926	0.981	0.990
15	S_N^*	0.958	0.989	0.994	0.999	1	1
	S_N	0.943	0.978	0.988	0.999	1	1
25	S_N^*	0.951	0.977	0.987	1	1	1
	S_N	0.977	0.992	0.996	1	1	1
$\sigma_{ij} = 0$							
5	S_N^*	0.620	0.827	0.906	0.958	0.994	0.997
	S_N	0.560	0.797	0.898	0.941	0.987	0.994
15	S_N^*	0.963	0.988	0.997	1	1	1
	S_N	0.977	0.998	1	1	1	1
25	S_N^*	0.959	0.981	0.985	1	1	1
	S_N	0.999	1	1	1	1	1

4.2 基于共同因素的数据生成过程

综列数据的截面相关可能来自某一共同因素的影响, 为了考察 GNIV 检验在这种情况下的表现, 我们借鉴 Pesaran 的数据生成过程, 生成具有截面相关的综列数据. 即依照模型 (1) 生成综列数据 $\{y_{it}\}$, 其中, $\{u_{it}\}$ 被设定为如下的过程:

$$u_{it} = \gamma_i f_t + \varsigma_{it} \tag{20}$$

$$\varsigma_{it} = \rho_i \varsigma_{i,t-1} + \varepsilon_{it} \tag{21}$$

其中 $\gamma_i \sim U[-1, 3]$, $f_t \sim iidN(0, 1)$ 为导致截面个体相关的共同因素, $\varepsilon_{it} \sim iidN(0, \sigma_i^2)$, $\sigma_i \sim U[0.5, 1.5]$. 在原假设下考察检验水平时, $\alpha_i = \alpha = 1$, ρ_i 被设定为 0.3, 在备选假设下考察检验势时, $\alpha_i \sim iidU[0.85, 0.95]$, $\rho_i \sim U[0.2, 0.4]$. 表 3 和表 4 分别给出了 GNIV 检验在有限样本下的检验水平和检验势、以及与 S_N 检验和 Pesaran 的 CIPS 检验的比较, 显著性水平分别选取为 0.01、0.05 和 0.10.

表 3 检验水平 (基于共同因素的截面相关设定, $\rho_i = 0.3, \alpha_i = 1$)

N		$T=50$ 1%	$T=50$ 5%	$T=50$ 10%	$T=100$ 1%	$T=100$ 5%	$T=100$ 10%
10	S_N^*	0.021	0.067	0.113	0.010	0.069	0.117
	S_N	0.007	0.048	0.090	0.000	0.037	0.073
	CIPS	0.003	0.039	0.087	0.025	0.083	0.147
15	S_N^*	0.017	0.064	0.095	0.025	0.076	0.114
	S_N	0.017	0.043	0.082	0.005	0.043	0.074
	CIPS	0.013	0.040	0.083	0.033	0.084	0.125
20	S_N^*	0.016	0.049	0.087	0.019	0.066	0.131
	S_N	0.006	0.049	0.092	0.006	0.031	0.062
	CIPS	0.001	0.013	0.062	0.024	0.068	0.105
30	S_N^*	0.008	0.019	0.033	0.010	0.044	0.076
	S_N	0.007	0.036	0.070	0.003	0.072	0.056
	CIPS	0.001	0.015	0.030	0.016	0.047	0.092

表 4 检验势 (基于共同因素的截面相关, $\rho_i \sim U[0.2, 0.4]$; $\alpha_i \sim U[0.85, 0.99]$)

N		$T=50$ 1%	$T=50$ 5%	$T=50$ 10%	$T=100$ 1%	$T=100$ 5%	$T=100$ 10%
10	S_N^*	0.913	0.972	0.988	1	1	1
	S_N	0.895	0.967	0.987	0.998	1	1
	CIPS	0.041	0.184	0.298	0.548	0.796	0.891
15	S_N^*	0.965	0.989	0.995	1	1	1
	S_N	0.974	0.995	0.998	1	1	1
	CIPS	0.060	0.201	0.307	0.734	0.907	0.954
20	S_N^*	0.978	0.0096	0.998	1	1	1
	S_N	0.993	1	1	1	1	1
	CIPS	0.048	0.168	0.276	0.789	0.916	0.964
30	S_N^*	0.961	0.985	0.990	1	1	1
	S_N	1	1	1	1	1	1
	CIPS	0.034	0.149	0.279	0.845	0.943	0.980

4.3 仿真结果分析

基于具体的仿真实验结果, 可以得到以下结论: ①一般性截面相关情形下, 我们的 GNIV 检验在各种截面相关程度下均具有良好的有限样本性质, 综列单位根设定下的拒绝概率非常接近相应的显著性水平. 而 Chang 的 S_N 检验则具有明显的分布扭曲, 并且截面相关性越强 (如相关系数 0.8), 扭曲也就越严重. GNIV 检验在各种设定情形下均具有较高的检验势, 并随着 N 和 T 的增大而提高. ②. 基于共同因素的截面相关情形下, 三种检验 (GNIV、CIPS、S_N) 均具有较好的有限样本性质, 综列单位根设定下的拒绝概率都比较接近对应的显著性水平. 只是在 $T=50$, $N=30$ 时, GNIV 检验和 CIPS 检验具有明显的分布扭曲. 在各种样本规模下, GNIV 和 S_N 检验的检验势都要高于 CIPS 检验.

总之, 仿真结果显示, GNIV 检验在各种设定情形中都具有良好的表现, 并且优于其他两种检验方法.

5 结论

本文在 Chang 的基础上, 将似无关回归和 Chang 的 NIV 估计结合起来, 通过即期协方差矩阵消除截面单元之间的相关性, 通过 NIV 估计得到极限分布为标准正态分布的各截面个体的 t 比率, 进而构建综列单位根的 GNIV 检验. 仿真实验结果显示, 我们所提出的 GNIV 检验无论在截面单元之间不相关、弱相关、还是强相关情形下, 无论是在一般性相关设定形式下, 还是在基于共同因素的相关设定形式下, 均具有良好的有限样本性质和较高的检验势, 明显优于 Chang 的 S_N 检验和 Pesaran 的 CIPS 检验. 所以说, 无论是对截面相关还是截面不相关的综列数据而言, GNIV 检验统计量都具有稳健性.

附录: 命题 1 的证明

因为 $\Delta y_{i,t-1},\cdots,\Delta y_{i,t-p}$ 都是平稳的, 从 (6) 可知, $\Delta y^*_{i,t-1},\cdots,\Delta y^*_{i,t-p}$ 也都是平稳的. 因为 $F(y^*_{i,t-1})$ 是规范可积的, 根据 Chang、Park 和 Phillips[8] 的命题 5(e)—— 单整过程的可积变换与平稳回归元之间是渐近正交的, 因而有

$$T^{-3/4}\sum_{t=1}^{T}F(y^*_{i,t-1})\Delta y^*_{i,t-k}\to_p 0,\quad k=1,\cdots,p$$

即:

$$\sum_{t=1}^{T}F(y^*_{i,t-1})X^*_i=o_p(T^{3/4}) \tag{A1}$$

显然, X^*_i 是 p 维的向量平稳过程, 根据 Chang 的推导过程:

$$\begin{aligned}&\left|\sum_{t=1}^{T}F(y^*_{i,t-1})X^{*'}_{it}\left(\sum_{t=1}^{T}X^*_{it}X^{*'}_{it}\right)^{-1}\sum_{t=1}^{T}X^*_{it}\varepsilon^*_{it}\right|\\&\le\left|\sum_{t=1}^{T}F(y^*_{i,t-1})X^{*'}_{it}\right|\left|\left(\sum_{t=1}^{T}X^*_{it}X^{*'}_{it}\right)^{-1}\right|\left|\sum_{t=1}^{T}X^*_{it}\varepsilon^*_{it}\right|\\&=o_p(T^{3/4})O_p(T^{-1})O_p(T^{1/2})=o_p(T^{1/4})\end{aligned} \tag{A2}$$

和

$$\begin{aligned}&\left|\sum_{t=1}^{T}F(y^*_{i,t-1})X^{*'}_{it}\left(\sum_{t=1}^{T}X^*_{it}X^{*'}_{it}\right)^{-1}\sum_{t=1}^{T}X^*_{it}F(y^*_{i,t-1})\right|\\&\le\left|\sum_{t=1}^{T}F(y^*_{i,t-1})X^{*'}_{it}\right|\left|\left(\sum_{t=1}^{T}X^*_{it}X^{*'}_{it}\right)^{-1}\right|\left|\sum_{t=1}^{T}X^*_{it}F(y^*_{i,t-1})\right|\\&=o_p(T^{3/4})O_p(T^{-1})o_p(T^{3/4})=o_p(T^{1/2})\end{aligned} \tag{A3}$$

定义

$$A_i = F(y^*_{i,-1})'\varepsilon^*_{it} - F(y^*_{i,-1})'X^*_i(X^{*'}_iX^*_i)^{-1}X^{*'}_i\varepsilon^*_{it}$$
$$B_i = F(y^*_{i,-1})'y^*_{i,-1} - F(y^*_{i,-1})'X^*_i(X^{*'}_iX^*_i)^{-1}X^{*'}_iy^*_{i,-1}$$
$$C_i = F(y^*_{i,-1})'F(y^*_{i,-1}) - F(y^*_{i,-1})'X^*_i(X^{*'}_iX^*_i)^{-1}X^{*'}_iF(y^*_{i,-1})$$

基于 (A2) 和 (A3), 可得:

$$T^{-1/4}A_i = T^{-1/4}\sum_{t=1}^{T}F(y^*_{i,t-1})\varepsilon^*_{it} + o_p(1) \tag{A4}$$

$$T^{-1/2}C_i = T^{-1/2}\sum_{t=1}^{T}F(y^*_{i,t-1})^2 + o_p(1) \tag{A5}$$

因而 (7) 中 $\hat{\alpha}_i$ 的 t 比率可表示为:

$$t_{\alpha_i} = \frac{B_i^{-1}A_i}{(B_i^{-2}C_i)^{1/2}} = \frac{T^{-1/4}\sum\limits_{t=1}^{T}F(y^*_{i,t-1})\varepsilon^*_{it}}{(T^{-1/2}\sum\limits_{t=1}^{T}F(y^*_{i,t-1})^2)^{1/2}} + o_p(1) \tag{A6}$$

根据 Chang、Park 和 Phillips[8] 的命题 5(c), 得到:

$$T^{-1/4}\sum_{t=1}^{T}F(y^*_{i,t-1})\varepsilon^*_{it} \to_d \left(L_i(1,0)\int_{-\infty}^{\infty}F(\pi_i(1)s)^2\mathrm{d}s\right)^{1/2}B(1) \tag{A7}$$

根据 Chang、Park 和 Phillips 的命题 5(i), 得到:

$$T^{-1/2}\sum_{t=1}^{T}F(y^*_{i,t-1})^2 \to_d L_i(1,0)\int_{-\infty}^{\infty}F(\pi_i(1)s)^2\mathrm{d}s, \tag{A8}$$

其中 $L_i(t,s) = \lim\limits_{\delta\to 0}\frac{1}{2\delta}\int_0^t 1\{|F(r)-s|<\delta\}\mathrm{d}r$, $\pi_i(1) = W^i(1)^{-1}$, W^i 的定义见式 (2), B(1) 表示标准布朗运动. 将 (A7) 和 (A8) 的结论代入 (A6), 则有:

$$t_{\alpha_i} \to_d \frac{\left(L_i(1,0)\int_{-\infty}^{\infty}F(\pi_i(1)s)^2\mathrm{d}s\right)^{1/2}B(1)}{\left(L_i(1,0)\int_{-\infty}^{\infty}F(\pi_i(1)s)^2\mathrm{d}s\right)^{1/2}} = B(1) \tag{A9}$$

同时, 对于 $i \neq j$, 当 $T\to\infty$, $\mathrm{cor}(\varepsilon^*_i,\varepsilon^*_j)\to_p 0$, 则 $\mathrm{cor}(y^*_{it},y^*_{jt})\to_p 0$, 进而有 $\mathrm{cor}(t_{\alpha_i},t_{\alpha_i})\to_p 0$. 因为 $\hat{\Sigma}$ 和 $\hat{\Gamma}$ 分别是 Σ 和 Γ 的一致估计, 故命题 1 成立.

参考文献

[1] Chang Y. Nonlinear IV unit root tests in panels with cross-sectional dependency[J]. Journal of Econometrics, 2002, 110(2): 261–292.

[2] Bai J, Ng S. A PANIC attack on unit roots and cointegration[J]. Econometrica, 2004, 72(4): 1127–1178.

[3] Choi I. Combination unit root tests for cross-sectional correlated panels[C]//Mimeo, HongKong University of Science and Technology, 2002.

[4] Im K S, Pesaran M H. On the panel unit root tests using nonlinear instrumental IV variables[C]//Mimeo, University of Southern California, 2003.

[5] Pesaran H M. A simple panel unit root test in the presence of cross section dependence[C]//Mimeo, University of Southern California, 2003.

[6] Elliott G, Rothenberg T, Stock J. Efficient tests for an autoregressive unit root[J]. Econometrica, 1996, 64(4): 813–836.

[7] So B S, Shin D W. Recursive mean adjustment in time series inferences[J]. Statistics & Probability Letters, 1999, 43(1): 65–73.

[8] Chang Y, Park J Y, Phillips P C B. Nonlinear econometric models with cointegrated and deterministically trending regressors[J]. Econometrics Journal, 2001, 4 (1): 1–36.